

CLAIMS**We Claim:**

1. A method for automatically extracting data from at least one electronic document including a target page being accessible over a computer network, said target page comprising a plurality of elements each having a structural definition, wherein said structural definitions interrelate said plurality of elements, said method comprising:

recording a sequence of actions operable to navigate to said target page of said electronic document;

automatically mapping said target page by identifying said structural definitions of said plurality of elements;

automatically identifying a target pattern for a selected subset of said plurality of elements based upon said structural definitions of said selected elements;

automatically accessing said target page according to said recorded sequence and remapping said target page by reidentifying said structural definitions of said plurality of elements; and,

automatically identifying and copying select ones of said plurality of elements dependent upon said target pattern and said reidentified structural definitions.

2. The method of Claim 1, wherein said electronic document is a website and said target page is a web page.

3. The method of Claim 2, wherein said actions include user interaction with said website.

4. The method of Claim 3, wherein said user interaction includes filling in forms and activating HTML links.
5. The method of Claim 1, further comprising:
altering said sequence of actions, thereby generating a second sequence of actions;
automatically executing said second sequence of actions thereby accessing an altered target page; and,
automatically identifying and copying select ones of said plurality of elements from said altered target page dependent upon said target pattern.
6. The method of Claim 1, further comprising:
altering said sequence of actions, thereby generating a second sequence of actions;
automatically executing said second sequence of actions thereby accessing a second target page; and,
automatically identifying and copying select ones of said plurality of elements from said second target page dependent upon said target pattern.
7. The method of Claim 1, further comprising auditing said method, said auditing comprising:
automatically accessing said target page according to said recorded sequence and remapping said target page by reidentifying said structural definitions of said plurality of elements;

comparing said select ones of said plurality of elements to predicted results to determine whether said method is functioning properly.

- generating at least one file including data being indicative of said actions and said target pattern; and,

9. The method of Claim 8, wherein said computer network includes a global interconnection of computer networks.

11. A method for automatically extracting data from a target page of at least one electronic document being accessible over a computer network, said target page comprising a plurality of elements each having a structural definition wherein said structural definitions interrelate said plurality of elements, said method comprising:

identifying a target pattern for a selected subset of said plurality of elements;

automatically accessing said target page according to said recorded sequence; and,
automatically identifying and copying select ones of said plurality of elements and
said structural definitions of said select ones of said plurality of elements dependent upon
said target pattern.

12. The method of Claim 11, wherein:

in a first mode, said target pattern is dependent upon said interrelation of said
structural definitions for said selected subset;

in a second mode, said target pattern is dependent upon content of said selected
subset; and,

in a third mode, said target pattern is dependent upon said structural definitions for
and content of said selected subset.

13. The method of Claim 12, wherein said actions include user interaction with a
plurality of electronic documents.

14. The method of Claim 13, wherein each of said plurality of electronic documents
comprises a web page.

15. The method of Claim 12, wherein said actions include activating HTTP links and
electronically filling and submitting forms.

16. A computerized system for automatically copying select data from a web site, data associated with said web site including a plurality of elements each having structural data associated with it, and being stored on a server being accessible over a computer network, said structural data and elements defining a select web page of said web site, said system comprising:

a navigation module being operable on a microprocessor based device electronically coupled to said computer network, said navigational module being operable to:

record a sequence of actions operable to electronically navigate to said select web page of said web site using said computer network; and,

automatically access said select web page according to said recorded sequence; and,

an extraction module being operable on said microprocessor based device, said extraction module being operable to:

identify a pattern for a selected subset of said plurality of elements on said select web page; and,

automatically identify and copy select ones of said plurality of elements and said structural definitions of said select web page dependent upon said pattern.

17. The system of Claim 16, wherein said microprocessor based device includes and said modules are operable through Internet browser software.

18. The system of Claim 17, wherein said modules and browser software are mutually responsive to one another.

19. The system of Claim 18, wherein:

in a first mode, said pattern is dependent upon said interrelation of said structural definitions for said selected subset;

in a second mode, said pattern is dependent upon content of said selected subset; and,

in a third mode, said pattern is dependent upon said structural definitions for and content of said selected subset.

20. The system of Claim 19, wherein:

said navigation module is adapted to automatically alter said sequence of actions according to predetermined criteria; and, automatically access a other web pages according to said altered sequence; and,

said extraction module is adapted to automatically alter said pattern according to predetermined criteria; and, automatically identify and copy other select ones of said plurality of elements and said structural definitions of said other web pages dependent upon said altered pattern.

009127 0205460